

二段階生成モデルによるスケッチベースの 錆テクスチャ生成手法の提案

舟橋 克樹[†] 謝 浩然[†] 宮田 一乗[†]

[†] 北陸先端科学技術大学院大学 〒923-1205 石川県能美市旭台 1-1

E-mail: [†] {s2110147, xie, miyata}@jaist.ac.jp

あらまし コンピュータで合成された画像は、写実性の観点においてテクスチャに大きく依存する。本研究では、実世界でよく見られる錆テクスチャに着目し、ユーザスケッチを入力とする錆テクスチャ生成システムを提案する。まず、錆テクスチャを制御マップに変換し、大規模変動や非局所的な特徴を抽出する。そして制御マップをスケッチデータに変換し、大局的な形状情報を抽出する。これにより、錆テクスチャから制御マップとスケッチを得る。この2つの画像と元の錆テクスチャから構築した学習データセットを用いて、ユーザスケッチから複雑な構造を持つ錆テクスチャを生成するための二段階の生成モデルを学習する。以上の提案手法を利用して、ユーザスケッチから錆テクスチャを生成するシステムを実装し、生成される錆テクスチャの入力スケッチとの忠実性、錆テクスチャとしての写実性、多様性について評価する。

キーワード 深層学習, 画像処理, 画像生成, CG, 錆テクスチャ

Sketch-Based Rust-Texture Design with Two-Stage Generative Model

Yoshiki Funahashi[†] Haoran Xie[†] and Kazunori Miyata[†]

[†] Japan Advanced Institute of Science and Technology 1-1 Asahidai, Nomi-shi, Ishikawa, 923-1205 Japan

E-mail: [†] {s2110147, xie, miyata}@jaist.ac.jp

Abstract Computer-generated images are highly dependent on texture in terms of realism. In this study, we focus on rust textures commonly found in the real world and propose a rust texture generation system using user sketches as input. First, the rust texture is converted into a control map to extract large-scale variations and nonlocal features. Then, the control map is converted to sketch data to extract global shape information. With the control map and sketch from the rust texture, we constructed the training dataset for texture image generation. A two-stage generative model is proposed to generate a rust texture with a complex structure from a user sketch. Using the proposed method, we implemented a prototype system that generates rust textures from user sketches, and evaluated the fidelity of the generated rust textures to the input sketches, their realism and diversity as rust textures.

Keyword Deep learning, Image processing, Image generation, CG, Rust-Texture

1. はじめに

近年、CGは急速に発展している。メタバースやゲームにおいて、オブジェクトのテクスチャの視覚的詳細の欠如から人工的な印象を与え、ユーザの没入感を妨害してしまうことがしばしばある。実世界に存在する全てのオブジェクトは、少なからず、汚れや傷といった非常に多くの視覚的な詳細を持った不均質なテクスチャで構成されている。メタバースやゲームで生成されるシーンをより写実的に表現するには、シーン内のオブジェクトに付随する不均質なテクスチャをリアルかつ自然に生成することが重要である。

リアルなテクスチャを得るための手法として、実世

界の物体や風景をスキャンしたり、写真を撮影したりすることであるが、これらの作業はコストがかかる作業である。また、コンピュータによって一から制作する方法も考えられるが、不均質テクスチャの制作は高い編集スキルが必要であり、非常に困難な作業である。

これらを受けて、よりリアルな不均質テクスチャを生成するための研究が行われている。近年では、所望のテクスチャの大規模な変動と非局所的な特徴を符号化することで不均質テクスチャの特徴を捉える制御マップによって、テクスチャ合成プロセスを導く手法が提案されている[1][2][3][4]。

しかし、制御マップの持つ元のテクスチャの大規模

な変動と非局所的な特徴情報を維持したまま、制御マップを編集する手法は管見の限りまだない。これらの問題を解決することで、所望の形状を持ち、かつ自然な不均質テクスチャの合成を素早く行うことが可能となる。これは、映像クリエイターや 3D CG の分野で、コンピュータで実世界を再現することが非常に容易になるという貢献が期待できる。

ここで、錆テクスチャは実世界に比較的多くみられる不均質テクスチャである。近年、意図的に錆テクスチャを生成することはインテリアのデザインや、プラモデルの塗装などの芸術分野でしばしば行われる。また、前述したメタバースやゲームのシーンを写実的、かつより印象的にするためにもしばしば使用される。このように芸術分野や CG 分野で広く扱われる錆テクスチャは、非常に複雑な不均質な構造を持つ。錆テクスチャの例を図 1 に示す。錆テクスチャは錆が発生している部分(図 1(a))、発生した錆による副産物である副錆により、元のテクスチャが変色している部分(図 1(b))、錆が発生していない元のテクスチャ部分(図 1(c))により構成されていることが観察できる。

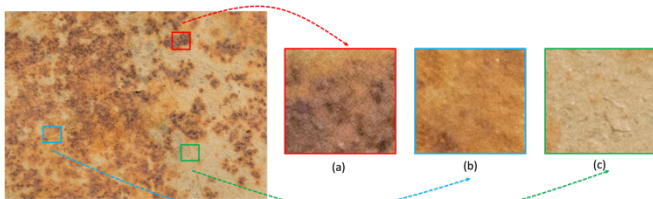


図 1: 錆テクスチャの例¹

また、最近ではスケッチを入力とした制作支援の研究が存在する[5][6][7][8]。スケッチは、大域的な形状や変動の流れなどの情報を把握しやすく、かつ加筆・修正が容易である点から、様々なデザインの足がかりとして適している。一方で、スケッチの持つ情報量は非常に少なく、曖昧であるため、詳細なデザインを行うことは困難である。

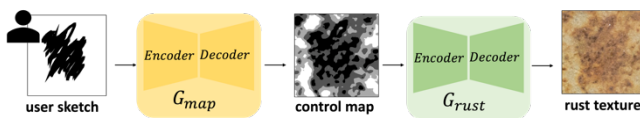


図 2: システム概要図

図 2 に本研究の概要図を示す。本研究では、前述したように芸術分野や CG 分野で広く扱われており、非常に不均質な構造を持つ錆テクスチャに注目し、ユーザスケッチを入力とした錆テクスチャ合成手法を提案する。スケッチから錆テクスチャを生成する過程を、

1)情報量の少ないスケッチから、錆テクスチャの錆による複雑な変色の広がりや錆度合いによってラベル化した制御マップを生成する。2)得られた制御マップから各ラベル化された領域に対応するテクスチャを合成することにより、錆テクスチャを生成する。という 2 段階の画像変換問題と捉えることで、大域的な形状や変動の流れなどの大まかな情報のみを持つスケッチから複雑な不均質構造を持つ錆テクスチャの生成を可能にする。 G_{map} , G_{rust} では、条件付き敵対的生成ネットワーク(以降、cGAN とする)を用いる。また、生成される錆テクスチャについて、スケッチとの忠実性、錆テクスチャとしての写実性、多様性の観点から考察し、さらにアブレーションスタディを実施することでのシステムの有効性を評価する。

2. 関連研究

2.1. 制御マップを利用した画像合成手法

制御マップを用いた画像合成手法の研究は多く行われている。Bellini ら[4]は、実世界で撮影された 1 枚のテクスチャに対して、繰り返される画像パッチは元のテクスチャパターンの一部である可能性が高く、繰り返されないパッチは風化作用の結果である可能性が高いという重要な観察結果を利用して、Age Map という制御マップを開発した。その後、Chen ら[9]によって、Isola ら[10]が提案した cGAN を利用した画像変換モデルである pix2pix モデルを用いることにより、Age Map からテクスチャ画像の合成の高速化に成功している。

Rosenberger ら[1]は、不均質である自然な質感のテクスチャの多くに適した制御マップは、複数の層の重ね合わせとしてモデル化でき、各層の可視部分はより均質なテクスチャによって占有されるという観察に基づいて、テクスチャを少数のレイヤーに分解した後、新しいレイヤーのセットを自動的に生成するための新しい形状合成アルゴリズムを開発した。

Lockerman ら[2]は、色やテクスチャが類似するピクセルで構成された小領域であるスーパーピクセルの階層的拡張を行い、次にスーパーピクセル間の類似性グラフにおけるランダムウォークと非負行列因子分解に基づいてスーパーピクセルにラベル付けを行うという 2 段階のプロセスによって、完全に自動化された制御マップの生成を可能にした。

2.2. スケッチ入力による制作支援

ユーザスケッチを入力として、画像や 3D オブジェクト、ガイダンスなどを生成する研究が行われている。Li ら[5]は、合成された画像に含まれる大域的な形状情報を探索・利用し、スケッチから写実性の高い合成画像を生成する新たなフレームワークを提案した。

¹ <https://www.flickr.com/photos/ioachimphotos/32807630943/in/album-72157680933947446/>

また, Xie ら[6]は, さまざまなパラメータを設定する必要があり, 一般ユーザにはデザインが困難な煙の流体シミュレーションに注目し, 2 段階の条件付き敵対的生成ネットワークを利用することで, ユーザスケッチから軽量に動作し, かつそれらしい煙の流体シミュレーションを可能にした.

2.3. 本研究の位置付け

本研究の位置付けについて述べる. これまで, 制御マップを利用した不均質テクスチャの合成手法が多く研究されている[1][2][3][4]が, 制御マップの持つ特徴情報を維持したまま, ユーザが容易に制御マップの編集を行う手法はまだ確立されていない. そこで, 近年取り組まれてきたスケッチ入力による制作支援の研究[5][6][7][8]を参考に, スケッチデータによる直感的な制御マップの編集を可能にする. これにより, 直感的にデザインできるスケッチから複雑な構造を持つ不均一テクスチャの生成を目指す.

3. 提案手法

3.1. 概要

本研究の目的は, 錆テクスチャの複雑な不均質構造に従った写実性の高い, かつユーザがスケッチによって指定した大域的な形状情報に従った錆テクスチャを合成することである. 本研究の提案手法を図 3 に示す.

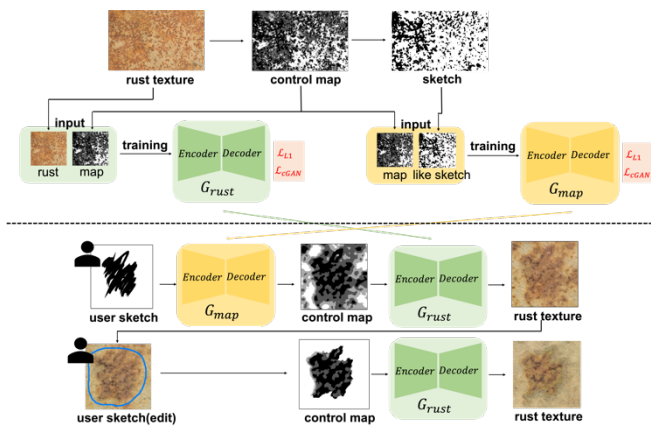


図 3: 提案手法

まず, 学習データの錆テクスチャから k-means++法[11]を用いて制御マップを生成する. 次にその制御マップに対して, 錆部分のマスク処理, B-スプライン曲線を利用した錆部分の輪郭線の簡略化の 2 つの処理によって, スケッチを生成する. ここで得られたスケッチ, 制御マップ, 錆テクスチャを学習データとして, スケッチと制御マップ, 制御マップと錆テクスチャとのそれぞれの関連性を Isola ら[10]が提案した cGAN を利用して学習する.

訓練された 2 つのネットワークモデルを用いることで, 錆テクスチャ生成プロセスでは, 錆が発生してい

る部分, 発生した錆による副産物である副錆により変色している部分, 錆が発生していない部分に至るまでの大規模な変動にのみ拘束されるため, ユーザにより入力されるスケッチの持つ大域的な形状情報に従った, 写実性の高い錆テクスチャが合成できる.

3.2. スケッチから制御マップの生成

錆テクスチャから, 制御マップを生成する手法について述べる. 制御マップの例を図 4 に示す.

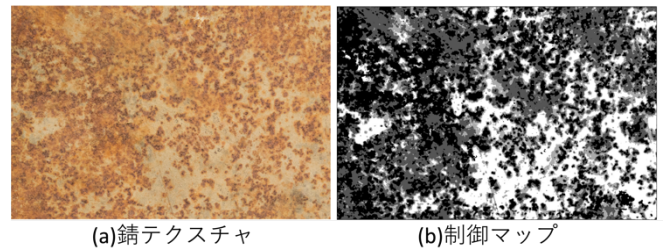


図 4: 制御マップの例

制御マップの生成で重要な点は, 入力された錆テクスチャの異なる特徴を持つそれぞれで均質なテクスチャ領域の持つ大規模な変動と非局所的な特徴を保持しつつ, それぞれの領域に異なるラベルを付与することで分類し, 情報量を落とすことである. 図 4 の例では, 図 4(a)で示すように, 元となる錆テクスチャ(図 4(b))の最も錆が発生している部分を黒く, 錆の発生度合いが弱い部分ほど白く表現することで, 錆の発生度合いの大規模な変動と, 各色で分割された錆度合いによって異なる領域の位置関係という非局所的な特徴を表現されていることが求められる.

錆テクスチャから制御マップを生成する研究は広く行われてきている. Bellini[4]らは, テクスチャの各異なる意味を持つ領域の認識には, 輝度値の他に, 明度の勾配方向も特徴量として利用することで, 認識精度が向上することを発見した. また Rosenberger[1]らは, 各画素の $N \times N$ 近傍の値を連結した N^2 次元の特徴ベクトルに対して k-means 法によりクラスタリングを行い, 錆テクスチャから離散値による制御マップを生成している.

本研究では, 入力する錆テクスチャの自由度が比較的確保される Rosenberger[1]らの手法に加えて, Bellini[4]らの発見による特徴量として明度の勾配方向を利用する. 学習データの錆テクスチャから, HSV 色空間の各要素に明度の勾配方向を加えた 4 つの特徴量に対して, k-means++法[11]により領域分割することで, 制御マップを生成する.

具体的には, まず, 入力テクスチャ画像を HSV 色空間に変換する. 次に各画素の明度の勾配方向を以下のようにして求める. 入力画像に対して, prewitt フィルタ処理をする. prewitt フィルタとは, 式(1)で示すカー

ネルによって、局所的な明度の変化を強調することができるものである。式(2)で示すように、入力画像 I と縦方向、横方向の prewitt フィルタの畳み込みによって得られる出力値 G_x, G_y は、それぞれの方向についての各画素の明度の変化量の大きさと捉えることができる。最終的に式(3)により明度の勾配方向を G 計算する。

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, K_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (1)$$

$$G_x = I * K_x, G_y = I * K_y \quad (2)$$

$$G = \tan^{-1} \frac{G_y}{G_x} \quad (3)$$

次に、HSV 色空間に変換した入力テクスチャの各画素値(h,s,v)に明度の勾配方向 G を加えた特徴量(h,s,v,G)を持つ画像に $m \times m$ (本研究では、 $m = 15$)の平均化フィルタを適用する。次に、得られた画像に対して k-means++法[11]を適用し、 k (本研究では $k = 4$)の領域に分割する。

最後に、得られた各領域に対して、錆度合いを表す色を、学習データの錆テクスチャと比較しながら設定していくことで、図 4 に示すような制御マップを得ることができる。本研究では、最も錆度合いが強い領域を黒、錆度合いが最も弱い領域を白としている。

3.3. 制御マップからスケッチデータの生成

3.2 節で得た制御マップからスケッチを生成する。スケッチの生成において重要な点は、制御マップによって表現された異なる特徴を持つテクスチャ領域の大規模な変動と非局所的な特徴情報を大域的な形状情報に単純化することによってスケッチに近い画像に変換する点である。本研究では、錆が発生している部分領域のみの形状をスケッチによって指定し、錆が発生している部分によって出現する変色領域の広がりや自動生成することに注目している。そのため、スケッチを生成する上で注目する領域は最も錆度合いが大きい領域、すなわち錆が発生している部分領域に限られる。また、本研究では、領域を塗りつぶすようなスケッチを採用し、スケッチの生成の際は、錆が発生している領域を大まかに塗り潰したような画像を生成する。スケッチ生成の各過程の処理結果を図 5 に示す。まず、制御マップ(図 5(a))から最も錆が発生している部分として表現される黒い領域以外を白で塗りつぶすことで、錆が発生している部分のみが黒で表された錆領域マスク画像を得る(図 5(b))。次に得られたマスク画像から黒く塗られた領域(錆が発生している部分領域)に対して、輪郭抽出処理を行う(図 5(c))。次に、輪郭線によって囲まれた領域の画素数が閾値 N 以下(本研究では、 $0 \leq N \leq 5000$)の領域を削除する(図 5(d))。

残った領域に対して輪郭線に沿った順序で輪郭点の抽出を行う。次に、抽出された輪郭点を輪郭線に沿って n (本研究では、 $n = 100$)ごとにサンプリングする(図 5(e))。この処理により、制御点の数が輪郭線に沿って均等に $1/n$ にダウンサンプリングされ、錆が発生している部分領域の単純化された輪郭線を得ることができる。ここで図 5(e)からわかるように、削減した輪郭点を結んだだけで、滑らかな輪郭線が得られないため、最後に削減して得られた輪郭点を制御点として B-スプライン補完処理をし、得られた滑らかな輪郭線の内側の領域(錆が発生している部分領域)を黒く塗りつぶし、スケッチを得る(図 5(f))。

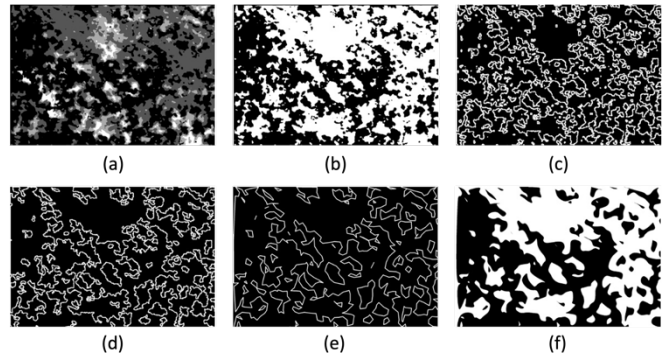


図 5: スケッチ化の流れ。それぞれ(a)制御マップ、(b)錆が発生している部分のマスク画像、(c)輪郭抽出処理後の画像、(d)小領域の輪郭線を削除した不画像、(e)輪郭点を均等にダウンサンプリングした画像、(f)スケッチを示す。

3.4. データセットの構築

本研究では、学習データ構築のための学習データの錆テクスチャを、写真の共有を目的としたコミュニティウェブサイトである Flickr からダウンロードすることで収集した。cGAN の生成結果を安定させるために、収集する錆テクスチャは、発生している錆の種類、錆が発生していない部分のテクスチャがある程度統一されている必要がある。実世界に存在する錆は、白系塗料を塗布された鉄素材に発生したものが多い。そこで本研究では、収集が比較的容易な白系塗装が施された表面に赤錆が進行している状態の錆テクスチャのみに限り、合計 10 枚の錆テクスチャを収集した。最終的な学習データは、スケッチ生成処理を行った上で得られた各画像を小領域に分割したものを利用するため、合計 700 枚で構築される。得られたスケッチ、制御マップ、錆テクスチャを図 6 に示す。

cGAN の学習に用いるスケッチデータ、制御マップ、錆テクスチャは、それぞれの画像全体を入力として直接使用するのではなく、画像を $M_{[pixel]} \times M_{[pixel]}$ の小領域に分割したパッチ(本研究では $M = 256$)をランダムに切り出すことで、データ数を拡張する。その後、生

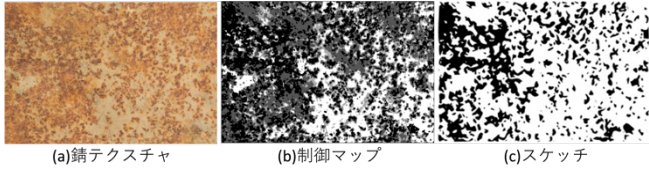


図 6: 得られたデータ例

成器のベースネットワークである U-Net の入力サイズに合わせて $128_{[pixel]} \times 128_{[pixel]}$ にダウンスケールする.

スケッチから制御マップを生成する cGAN の学習データは複数構築する. これは, ユーザに対して, 出力される錆テクスチャの条件を選択できる自由度を確保することを目的として, スケッチから制御マップの生成器を複数生成するためである. 3.3 節で述べた制御マップからスケッチの生成手法において, 図 5(d) で示される小領域の輪郭線を削除した画像を得るための閾値 N を, $0 \leq N \leq 5000$ の範囲で 1000 毎変更することで, 最終的に計 6 種類のスケッチを生成する. 各閾値 N によって得られた cGAN の学習データ用に切り出し, リサイズされたスケッチの例を図 7 に示す.

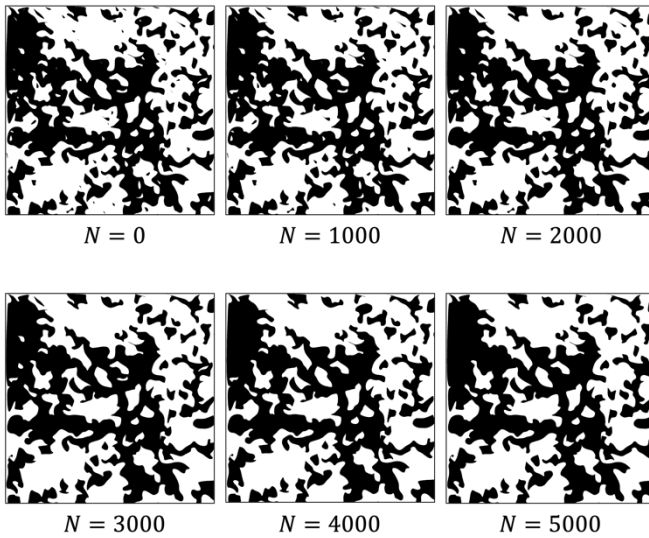


図 7: スケッチ例

3.5. 生成器と識別器の学習

構築されたデータセットをもとにスケッチから制御マップ, 制御マップから錆テクスチャを生成する 2 つの生成器の学習を行う. ネットワーク構造を図 8 に示す.

学習データとして, スケッチと制御マップ間の生成器ではスケッチと制御マップのペアデータ, 制御マップと錆テクスチャ間の生成器では制御マップと錆テクスチャのペアデータを利用する. 本研究では, スケッチデータ s からの制御マップ生成器 G_{map} と, 生成された制御マップ m からの錆テクスチャ生成器 G_{rust} に Isola[10]らにより提案された cGAN ベースの生成ネッ

トワークを用いている. また, 識別器 D_{map} と, D_{rust} は PatchGAN と同様に構成されている. 制御マップ生成モデルの損失関数はクロスエントロピーを利用して設計されており, 式(4)で表現される.

$$\min_{G_{map}} \max_{D_{map}} \mathbb{E}_{(s,m)} [\log D_{map}(s,m)] + \mathbb{E}_s [\log (1 - D_{map}(s, G_{map}(s)))] \quad (4)$$

ここで, $\mathbb{E}_{(s,m)}$ と \mathbb{E}_s は, それぞれスケッチデータ s と制御マップ m の生成モデルに対する期待値を表す.

錆テクスチャ生成モデルの損失関数は制御マップ生成モデルの損失関数と同様に式(5)で表される.

$$\min_{G_{rust}} \max_{D_{rust}} \mathbb{E}_{(m,r)} [\log D_{rust}(m,r)] + \mathbb{E}_m [\log (1 - D_{rust}(m, G_{rust}(m)))] \quad (5)$$

ここで, $\mathbb{E}_{(m,r)}$ と \mathbb{E}_m は, それぞれ制御マップ m と錆テクスチャ r の生成モデルに対する期待値を表す.

これらの生成モデルは, バッチサイズ 64, エポック数 200 で学習させた. 使用した最適化アルゴリズムは, 生成器と識別器の両方に Adam を使用し, 学習率は 0.0002 である.

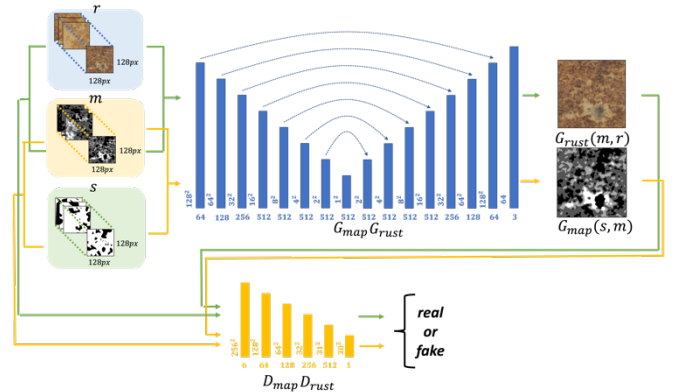


図 8: 生成器と識別器のネットワークモデル

4. システム実装

本システムのフレームワークを図 9 に示す.

4.1. スケッチデータからテクスチャの生成

ここでは, 3.5 節で学習した cGAN を利用して, ユーザスケッチから制御マップを生成し, 得られた制御マップから錆テクスチャを生成する. まず, ユーザ入力によって得られたスケッチから, 3.5 節で学習した G_{map} を利用して, 制御マップを生成する. ここで, G_{map} は学習段階で, $M_{[pixel]} \times M_{[pixel]}$ の小領域に分割されたパッチを学習データとしているため, ここで得られたスケッチも同様に各画像を, $M_{[pixel]} \times M_{[pixel]}$ (本研究では, $M = 256$) のパッチに分割処理をした後に, G_{map} に入力する. また, スケッチから制御マップを生成する際は, ユーザが生成される制御マップに対してスケッチによって表現された大域的な形状との忠実さと錆テクスチ

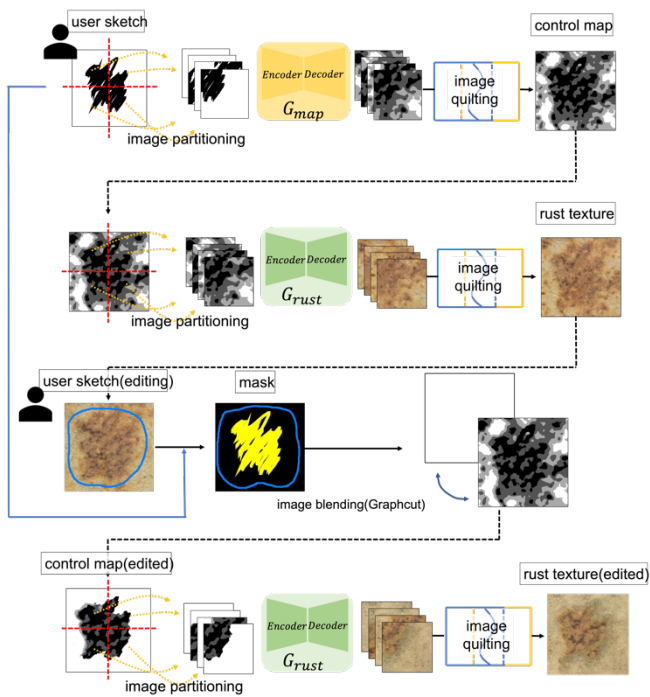


図 9: システムフレームワーク

そのものの写実性のトレードオフに何を求めるかを選択できるようにする. そのため, 3.4 節で生成した複数の学習データによって学習された複数の G_{map} 全てを利用して, 複数の制御マップを生成する.

制御マップから錆テクスチャの生成においては, 3.5 節で学習した G_{rust} を利用する. 基本的な生成プロセスは, スケッチから制御マップを生成するプロセスと同様である. 制御マップから錆テクスチャを生成する際は, スケッチから制御マップの生成段階で得られた複数の制御マップに対して, 錆テクスチャを生成する.

cGAN によって得られた制御マップと錆テクスチャは, $M_{[pixel]} \times M_{[pixel]}$ のパッチに分割された画像である. したがって, 最終的な制御マップと錆テクスチャを得るためには, 分割された画像をタイル状に並べ, 各画像の境界を自然に合成することで 1 枚の画像にする必要がある. 本研究では, 境界合成手法として, Efros ら [12] が提案した Image quilting アルゴリズムを利用する.

4.2. テクスチャの再編集

生成した錆テクスチャの再編集機能の実装について述べる. 本システムで学習した cGAN は実世界で撮影された錆テクスチャ画像から錆テクスチャの自然な発生, 広がり进行学习しているため, 本システムによって出力される錆テクスチャは, スケッチによって表現された領域以外の領域にも少量の錆が発生する. これは, 実世界に存在する錆テクスチャとしての写実性の観点では肯定される仕様だが, スケッチで表現された大域的な形状との忠実さの観点では, 否定的な仕様である.

本システムでは, 出力結果の錆テクスチャに対して, 上記で述べた, スケッチで表現された大域的な形状との忠実さと, 実世界に存在する錆テクスチャとしての写実性との間のトレードオフの選択機能として, テクスチャの再編集機能を実装した.

テクスチャの再編集を可能にする提案手法は, システムによって生成された元の制御マップと, 錆の発生が抑制されている制御マップ(本システムの制御マップでは, 錆が発生していない領域は白に近い色で表現されるため, 白に近い領域でのみ構成された制御マップ. 以降, 抑制制御マップとする)を合成するというものである. 元の制御マップの錆の生成を望まない領域に抑制制御マップを合成することによって所望の錆テクスチャの元となる制御マップが再生成される. 図 10 に制御マップの合成手法の概要を示す.

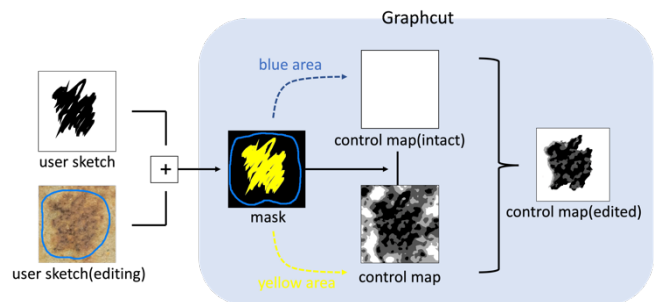


図 10: 制御マップの再生成手法

まず, 出力結果の錆テクスチャに対して, 錆の発生を望まない領域を表現したスケッチが入力されると, システムの序盤で入力されたスケッチと合成され, 新たな制御マップを生成するためのマスク画像が生成される. このマスク画像は, 錆の発生を望まない領域を表現したスケッチ(水色)とシステムの序盤で入力されたスケッチ(黄色)の中間領域がそれぞれの制御マップの境界となるおおよその位置を示す. ここで重要な点は, マスク画像で表現された領域内にそれぞれの制御マップが自然に重なり合うような境界線を探ることである.

ここでの境界の探索にあたって, 本研究では Kwatra ら [13] が提案した手法を利用する. この手法は, 既存の境界線探索アルゴリズムである Image quilting をベースに, 境界領域の動的計画法による最短経路問題に対して Graphcut を利用した最短経路問題に置換したアルゴリズムである.

5. 結果と考察

ここでは, 様々な種類のスケッチを用いて本システムによって生成した錆テクスチャの例を観察し, 本研究の提案手法の評価を行う. また, 本研究の提案手法の有効性をスケッチで表現された大域的な形状との忠

実さと、実世界に存在する錆テクスチャとしての写実性(以降、忠実性と写実性とする)、生成される錆テクスチャの多様性の観点から実証する。5.1 節で忠実性と写実性について考察し、5.2 節で、多様性について考察する。図 11 に本システムにより生成された錆テクスチャの例を示す。

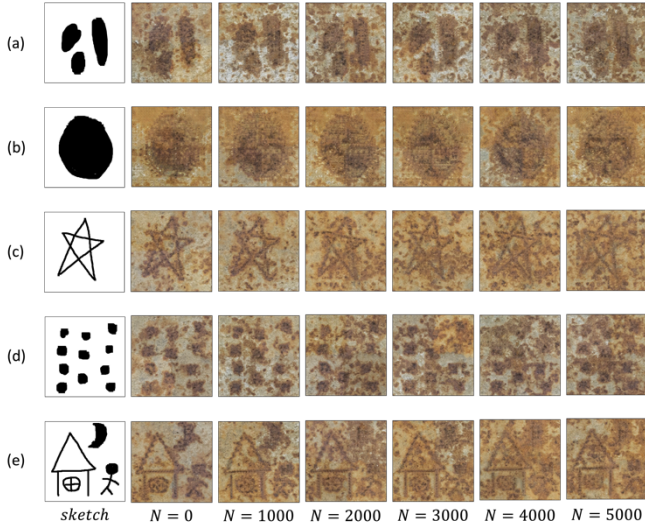


図 11: 本システムにより生成された錆テクスチャ

左端が入力されたスケッチを示し、右側に本システムにより生成した錆テクスチャを示す。本システムでは、3.3 節で述べた制御マップからスケッチの生成における閾値 N の違いによって、複数の錆テクスチャが生成される。 N は、制御マップからスケッチの生成手法において、図 5(d) で示される小領域の輪郭線を削除した画像を得るための閾値 N である。

5.1. 忠実性と写実性

図 11 から、 N が大きいほど、錆の広がり方が自然であり、錆の自然な広がりの影響で、スケッチで表現された大域的な形状が認識しにくくなっていることがわかる。一方で、 N が小さいほど、錆が発生している部分の周辺にほとんど錆の広がりによる変色が見られず、スケッチで表された大域的な形状が強く表現されていることがわかる。

図 12 に錆テクスチャの再編集の結果を示す。水色のスケッチによって指定された領域の錆を十分に除去できていることが確認できる。しかし、再生成された錆テクスチャは、錆が除去された部分の影響を受け、錆度合いによる錆テクスチャの見え方の違いが滑らかになってしまい、全体的にメリハリのないボヤけた錆テクスチャになってしまっている。これは、錆の除去により再生成された制御マップにおいて、錆度合いの変動が小さくなることによって、再生成された制御マップから錆テクスチャの生成で、メリハリのないボヤけた錆テクスチャが生成されてしまっていると考えら

れる。

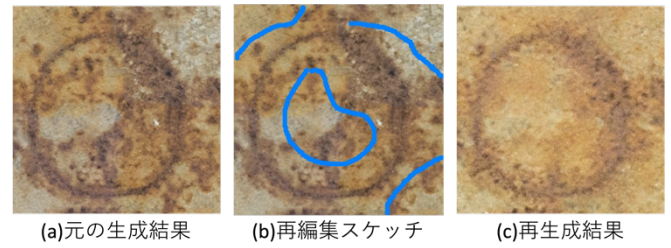


図 12: 錆テクスチャの再編集結果

5.2. 多様性

図 11 から様々な構成のスケッチに対応して、もっともらしい錆テクスチャが生成できることが確認できる。しかし、テクスチャの一部が不自然に合成されている例も確認できる。図 11(d), $N = 3000$ の画像を見ると、右上の領域だけ明るくなってしまっていることがわかる。これは、4.1 節で述べた入力されたユーザースケッチを $M_{[pixel]} \times M_{[pixel]}$ の小領域パッチに分割し、各領域で個別に生成処理を実施していることが考えられる。また、学習データの生成時の錆テクスチャから制御マップを生成する際に、異なる錆テクスチャにおいて、制御マップ生成のアルゴリズムが統一されていない点に関して原因の一つであると考えられる。

本研究では、錆テクスチャから制御マップへの変換は、各錆テクスチャで個別に k -means++ によって領域分割することでおこなっているため、画像内での特徴量のみ依存してしまい、制御マップの生成アルゴリズムが統一されていないという問題が起こる。異なるアルゴリズムで生成された各制御マップのデータを混在させて一つの学習データとしてしまっているため、このデータによって構築される cGAN の出力結果が不安定になってしまい、図 11(d), $N = 3000$ のように不自然な錆テクスチャが合成されてしまっていると考えられる。この問題を解決するためには、制御マップ生成手法を異なる画像間で統一の色変換規則となるように改良する必要がある。

5.3. アブレーションスタディ

本研究ではスケッチから錆テクスチャを生成する手法において、スケッチから直接錆テクスチャに変換せず、間に制御マップへの変換を挟む 2 段階画像変換問題と捉えている。制御マップで、スケッチで表現することができない錆テクスチャの特徴を再定義することで、より写実性の高い錆テクスチャを生成することができる。錆テクスチャは錆が発生している部分、発生した錆による副産物である副錆により、元のテクスチャが変色している部分、錆が発生していない元のテクスチャ部分により構成されている。さらにこれらの構造が不均一に分布している。制御マップへの変換で、これらの複雑な特徴を持つ各部分領域の大規模な変動

を理解することで、スケッチで表現された錆が発生している部分の大域的な形状に従うだけでなく、より写実性の高い錆テクスチャを得ることができる。

cGAN を 1 つだけ利用して、スケッチから直接錆テクスチャ生成した結果と、本手法による cGAN を 2 つ利用して、スケッチから制御マップ、制御マップから錆テクスチャを生成した結果を図 13 に示す。

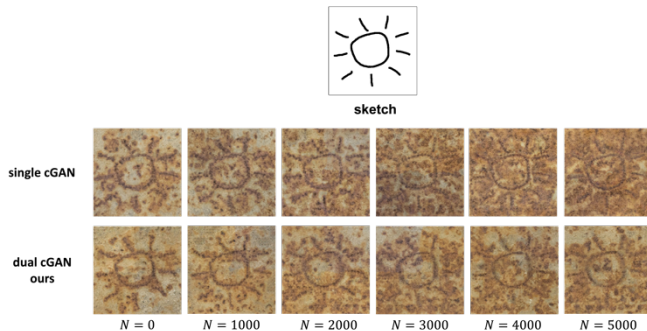


図 13: 一段階生成モデルと本手法の比較

どちらの結果も、強力な生成モデルである cGAN によって、最もらしい錆テクスチャが生成できている。しかし、cGAN を 1 つだけ利用した生成結果(図 13(上))は、スケッチによって表現された部分に均等に錆テクスチャが配置されているだけに過ぎず、実世界に存在する錆テクスチャのような不均一な分布が再現できていない。また、錆テクスチャの特徴である、錆が発生している部分が密集している領域は、変色している領域も大きく広がって分布するという特徴も再現できておらず、どの領域の変色部分も同じ大きさの広がりになってしまっている。一方、本手法に従って、cGAN を 2 つ利用した生成結果(図 13(下))は、大域的な形状は、スケッチに従っているが、それを基準に錆テクスチャが不均一に分布し、写実性の高い生成結果が得られていることがわかる。また、変色領域についても、錆が発生している部分が密集している領域は大きく、濃く広がっており、密集していない領域は小さく、薄く収まっていることが確認できる。これらの生成結果の違いは、 N の値が大きくなるにつれ、顕著に現れている。

6. 終わりに

本研究では、スケッチベースの二段階生成モデルを利用した錆テクスチャ生成システムを提案した。実世界に存在する錆テクスチャから大規模な変動と非局所的な特徴を抽出した制御マップに変換し、制御マップから大域的な形状情報を抽出したスケッチデータに変換することで、ユーザスケッチから複雑な構造を持つ錆テクスチャの生成が可能となった。

今後は、本システムを錆テクスチャに限らず、複雑な構造を持つ様々な不均一テクスチャに対しても利用可能なように各生成モデルを拡張し一般化すること、

もしくは、様々な不均一テクスチャそれぞれにおいて学習データを構築し、大規模な不均一テクスチャの学習データセットを構築することで、異なる不均一テクスチャそれぞれに対応した複数の生成モデルを学習することが求められる。

文 献

- [1] Rosenberger, Amir and Cohen-Or, Daniel and Lischinski, Dani. Layered shape synthesis: automatic generation of control maps for non-stationary textures, ACM Transactions on Graphics (TOG), Vol.28, No.5, pp.1-9, 2009.
- [2] Lockerman, Yitzchak David, et al. Multi-scale label-map extraction for texture synthesis, ACM Transactions on Graphics, 2016.
- [3] Lin, Yung-Chen and Way, Der-Lor and Shih, Zen-Chung. The synthesis of weathering using multi-label map, International Workshop on Advanced Imaging Technology (IWAIT) 2021, Vol.11766, pp.210-215, 2021.
- [4] Bellini, Rachele and Kleiman, Yanir and Cohen-Or, Daniel. Time-varying weathering in texture space, ACM Transactions on Graphics (TOG), Vol.35, No.4, pp.1-11, 2016.
- [5] Li, Zeyu and Deng, Cheng and Wei, Kun and Liu, Wei and Tao, Dacheng. Learning semantic priors for texture-realistic sketch-to-image synthesis, Neurocomputing, Vol.464, pp.130-140, 2021.
- [6] Xie, Haoran and Arihara, Keisuke and Sato, Syuhei and Miyata, Kazunori. DualSmoke: Sketch-Based Smoke Illustration Design with Two-Stage Generative Model, arXiv preprint arXiv:2208.10906, 2022.
- [7] Peng, Yichen and Mishima, Yuki and Igarashi, Yamato and Miyachi, Ryoma and Okawa, Masahiro and Xie, Haoran and Miyata, Kazunori. Sketch2domino: Interactive chain reaction design and guidance, 2020 Nicograph International (NicoInt), pp.32-38, 2020.
- [8] Funahashi, Yoshiki and Omori, Wataru and Yoshida, Shogo and Peng, Yichen and Xie, Haoran and Okada, Shogo and Miyata, Kazunori. Sketch2Makizushi: Interactive Guidance System for Decorative Sushi Rolls with User Sketches, 2022 Nicograph International (NicoInt), pp.44-47, 2022.
- [9] Chen, Li-Yu and Shen, I-Chao and Chen, Bing-Yu. Guided Image Weathering using Image-to-Image Translation, SIGGRAPH Asia 2021 Technical Communications, pp.1-4, 2021.
- [10] Isola, Phillip and Zhu, Jun-Yan and Zhou, Tinghui and Efros, Alexei A. Image-to-image translation with conditional adversarial networks, Proceedings of the IEEE conference on computer vision and pattern recognition, pp.1125-1134, 2017.
- [11] Arthur, David and Vassilvitskii, Sergei. k-means++: The advantages of careful seeding, 2006.
- [12] Efros, Alexei A and Freeman, William T. Image quilting for texture synthesis and transfer, Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp.341-346, 2001.
- [13] Kwatra, Vivek and Schödl, Arno and Essa, Irfan and Turk, Greg and Bobick, Aaron. Graphcut textures: Image and video synthesis using graph cuts, Acm transactions on graphics (tog), Vol.22, No.3, pp.277-286, 2003.